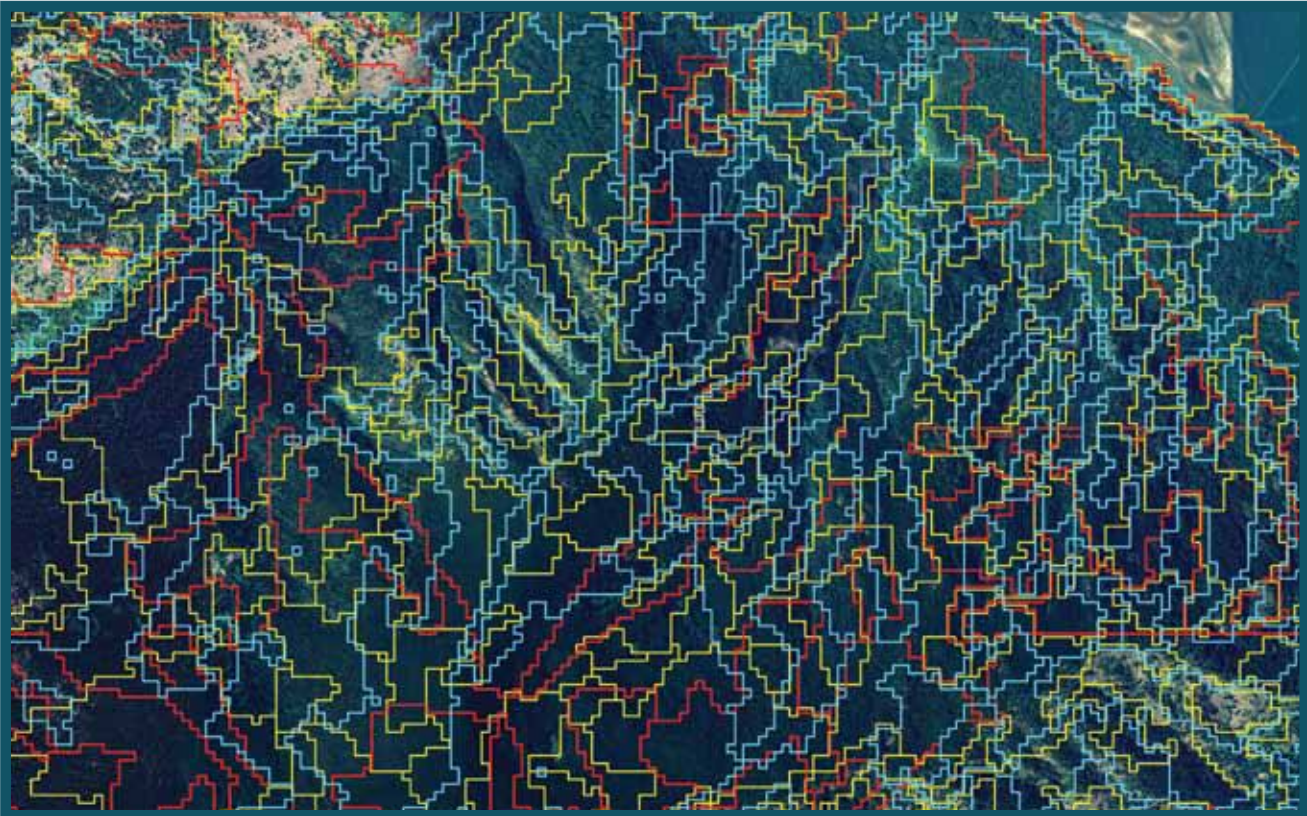


Toward Gleasonian Landscape Ecology: From Communities to Species, From Patches to Pixels

Samuel A. Cushman, Jeffrey S. Evans,
Kevin McGarigal, Joseph M. Kiesecker



United States Department of Agriculture / Forest Service

Rocky Mountain Research Station

Research Paper RMRS-RP-84

December 2010

Cushman, Samuel A.; Evans, Jeffrey S.; McGarigal, Kevin; Kiesecker, Joseph M. 2010. **Toward Gleasonian Landscape Ecology: From Communities to Species, From Patches to Pixels.** Res. Pap. RMRS-RP-84. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 12 p.

ABSTRACT

The fusion of individualistic community ecology with the Hutchinsonian niche concept enabled a broad integration of ecological theory, spanning all the way from the niche characteristics of individual species, to the composition, structure, and dynamics of ecological communities. Landscape ecology has been variously described as the study of the structure, function, and management of large heterogeneous land areas. Any reading of the published landscape ecology literature shows near uniformity in the adoption of a categorical patch-mosaic paradigm. However, if biological communities are multivariate gradients of species composition, with each species responding individually to particular combinations of limiting factors, is a categorical patch-based representation appropriate? We evaluate the sufficiency of several patch-based representations of vegetation at the landscape level to explain the composition of the plant community. Classified vegetation maps all performed poorly in explaining the composition and structure of forest trees among plots. Different categorical vegetation maps provided largely independent explanations of species variability. Individual species models based on spectral, topographic, and climatic variables vastly out-performed those produced using the classified maps. By moving from a landscape ecological paradigm based on categorical patches to one based on quantitative species and environmental responses across continuous space, it will be possible to both produce much more effective predictions of species distributions and ecological processes and remove much of the disjunction between landscape ecology and mainstream community ecology theory.

Keywords: Vegetation, variance partitioning, categorical maps, gradient theory, landscape ecology

THE AUTHORS

Samuel A. Cushman is a Research Ecologist with the USDA Forest Service, Rocky Mountain Research Station, in Missoula, Montana.

Jeffrey S. Evans is a Senior Landscape Ecologist with The Nature Conservancy—Central Science, in Fort Collins, Colorado.

Kevin McGarigal is a Professor with the Department of Natural Resources Conservation, University of Massachusetts, in Amherst, Massachusetts.

Joseph M. Kiesecker is a Director of Science with The Nature Conservancy—Central Science, in Fort Collins, Colorado.

You may order additional copies of this publication by sending your mailing information in label form through one of the following media. Please specify the publication title and series number.

Publishing Services

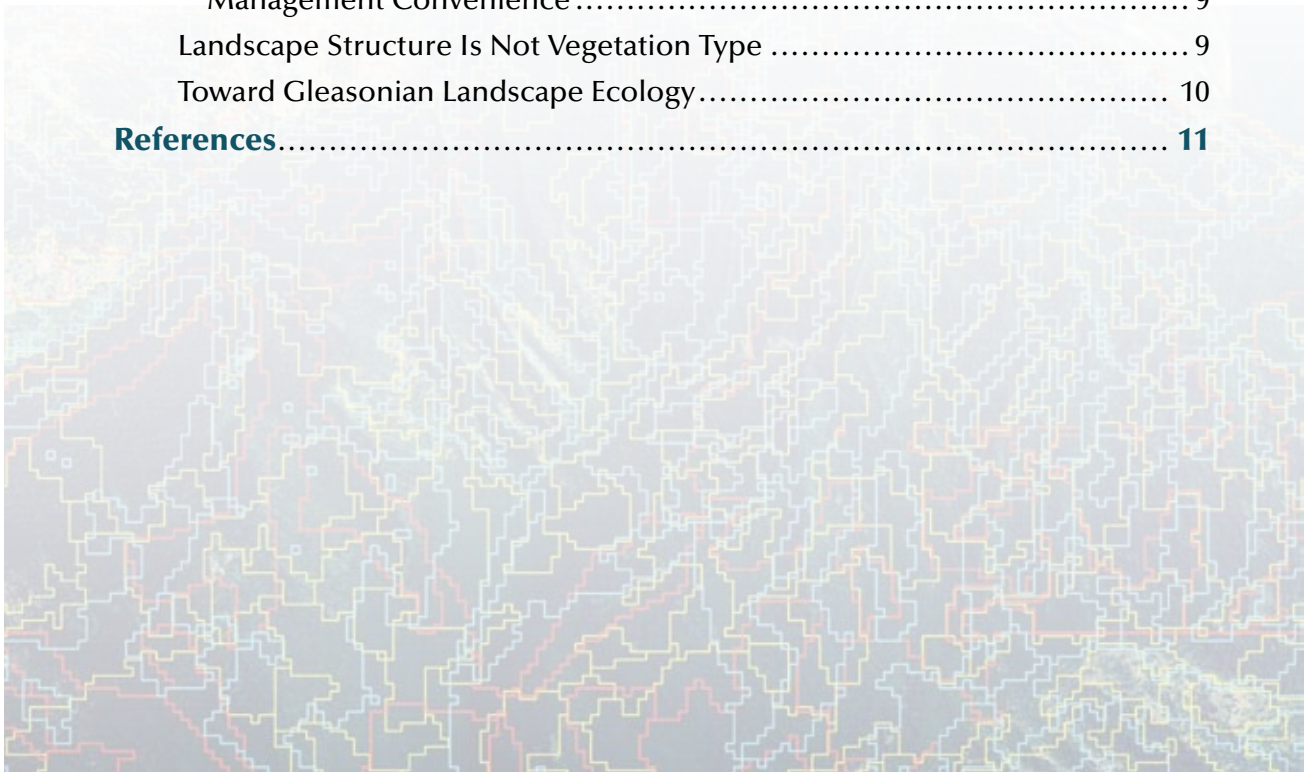
Telephone	(970) 498-1392
FAX	(970) 498-1122
E-mail	rschneider@fs.fed.us
Web site	http://www.fs.fed.us/rm/publications
Mailing address	Publications Distribution Rocky Mountain Research Station 240 West Prospect Road Fort Collins, CO 80526

Toward Gleasonian Landscape Ecology:

From Communities to Species, From Patches to Pixels

CONTENTS

Introduction	1
Clementsian Landscape Ecology	1
Objectives and Hypotheses	1
Methods.....	2
Vegetation Data.....	2
Classified Vegetation Maps	3
Independent Variables.....	4
Hypothesis 1: Variance Partitioning With Redundancy Analysis	4
Hypothesis 2: Predicting Occurrence of Four Major Species With Random Forest.....	4
Results	5
Hypothesis 1: Variance Partitioning With Canonical Correspondence Analysis.....	5
Hypothesis 2: Predicting Species Occurrence With Random Forests.....	5
Discussion	7
Hypothesis 1	7
Hypothesis 2	9
Categorical Landscape Ecology as Historical Legacy and Management Convenience	9
Landscape Structure Is Not Vegetation Type	9
Toward Gleasonian Landscape Ecology	10
References.....	11



Introduction

The modern scientific study of ecological communities is often traced to Clements (1907, 1916), who posited that the composition of species within a community is a deterministic product of regional climate and time since disturbance. Species composition within a community was thought to be highly predictable as deterministic functions of regional climate and seral condition. However, the Clementsian view of communities as analogs to super-organisms was fundamentally challenged by Gleason (1917, 1926), who argued that identification of categorical vegetation types was inconsistent with the heterogeneous structure of plant communities. He questioned delineating patch-mosaic maps of community types and opposed grouping of species in nameable associations. As an alternative, Gleason offered the individualistic concept of the plant association in which “the phenomena of vegetation depend completely upon the phenomena of the individual” species (Gleason 1917).

This individualistic concept of vegetation ecology is the foundation of modern community ecology. The fusion of individualistic community ecology (Gleason 1926; Curtis and McIntosh 1951; Whittaker 1967) with the Hutchinsonian niche concept (Hutchinson 1957) enabled a broad integration of ecological theory, spanning all the way from Darwinian evolution, to the niche characteristics of individual species, to the composition, structure, and dynamics of ecological communities. Each species is seen responding to local environmental and biotic conditions. The biotic community in this context is conceived as a collection of species that occur together at a particular place and a particular time due to overlapping tolerances of environmental conditions, biotic interactions (Lortie and others 2004), and vagaries of history, rather than an integrated and deterministic mixture. Research in this paradigm focuses on extending the individual concept to quantitative analysis of species distribution along environmental gradients and the effort to quantify the fundamental niche of each species in terms of the range of resources and conditions needed for that species to survive. The natural level of focus of such analyses is the species, not community type, assemblage, or patch type; the natural focal scale for such analyses is the location or pixel, rather than the stand or patch (McGarigal and Cushman 2005; Cushman and others 2007; Cushman and others 2010a).

Clementsian Landscape Ecology

Landscape ecology has been variously described as the study of the structure, function, and management of large heterogeneous land areas (Forman 1995) or, more generally, the study of spatial pattern and process (Turner 1989, 2005). Likewise, landscapes are typically described in terms of patches, corridors, and matrix (Forman 1995). These definitions explicitly frame the scope of landscape ecology within a patch-mosaic paradigm. This is a fundamentally Clementsian approach, and any reading of the published landscape ecology literature shows near uniformity in the adoption of this approach (McGarigal and others 2009). In the patch-mosaic paradigm (Forman 1995), each patch is implicitly treated like an individual of the super organism of each “patch type,” and it is assumed that by measuring the area and configuration of patch types we can represent the most important attributes of the landscape, including the distribution and abundance of plant and animal species.

However, if biological communities are multivariate gradients of species composition, with each species responding individually to particular combinations of limiting factors, is a categorical, patch-based representation appropriate? Put another way, isn't representing biological communities as categorical patches in a mosaic a de facto ratification of a Clementsian model of community composition at the landscape level? This disjunction between gradient-based, individualistic community ecology theory and patch-based, categorical landscape ecology has provoked some recent discussion (McIntyre and Barrett 1992; Manning and others 2004; McGarigal and Cushman 2005; Cushman and others 2008; McGarigal and others 2009; Cushman and others 2010c). However, these papers have largely been discussions of conceptual arguments and have had limited ability to provide real empirical evaluations of the strengths and limitations of alternative approaches.

Objectives and Hypotheses

Our goal in this paper is to evaluate the sufficiency of several patch-based representations of vegetation at the landscape level to explain the composition of the plant community. Our basic premise is that if a patch-based model of vegetation at the landscape-level actually represents the dominant scales and patterns of variation in the plant community, then classified vegetation

maps should explain a large portion of the variability in the density, basal area, and importance value of the tree species that comprise the community. The failure of classified vegetation maps to explain the structure and composition of the forest community would indicate that other approaches, such as species-based gradient approaches at the pixel scale, are needed to explain the patterns of biological communities at the landscape-level.

We break this central question into two hypotheses.

Hypothesis 1: If the structure and composition of the plant community is well represented by classified vegetation maps, which purportedly represent the major tree assemblages in this landscape, then these maps should explain (a) a large and (b) a similar amount of the variance in the variation of tree species abundance among plots. If these maps do not explain a large proportion of the variance in species abundance among plots then they do not represent the actual scales and patterns at which vegetation is primarily varying. Second, if different versions of these maps do not explain a similar portion of variance in the species assemblage then they do not provide a consistent measure of the vegetation community. As these maps were created with the same purpose, to describe the major patterns of the plant community, if they describe different portions of the structure of the species assemblage then this is an indication that these maps provide equivocal measures of the structure of the plant community.

Hypothesis 2: If the structure and composition of the plant community is well represented by classified vegetation maps, then those maps should be able to predict the occurrence of individual species among plots nearly as well as species-level, pixel-scale gradient models developed using spectral, topographical, and climatic variables. The combination of community-level (Hypothesis 1) and species-level (Hypothesis 2) tests provides the means to evaluate both classified maps' abilities to describe major patterns of co-occurrence and to predict the distribution of individual species.

Methods

Vegetation Data

The tree data used in this study were collected from a uniform grid of 411 vegetation plots distributed on a 1.6-km spacing across approximately 2500 square kilometers of northern Idaho, USA (Figure 1). Each plot consisted of a central 0.10 ha (17.95 m radius) plot and three 0.01 ha (7.32 m radius) subplots. For each plot the species and diameter were recorded for each tree, along with canopy closure, tree height, coarse woody debris, shrubs, forbs, and grasses. This analysis uses the tree data only.

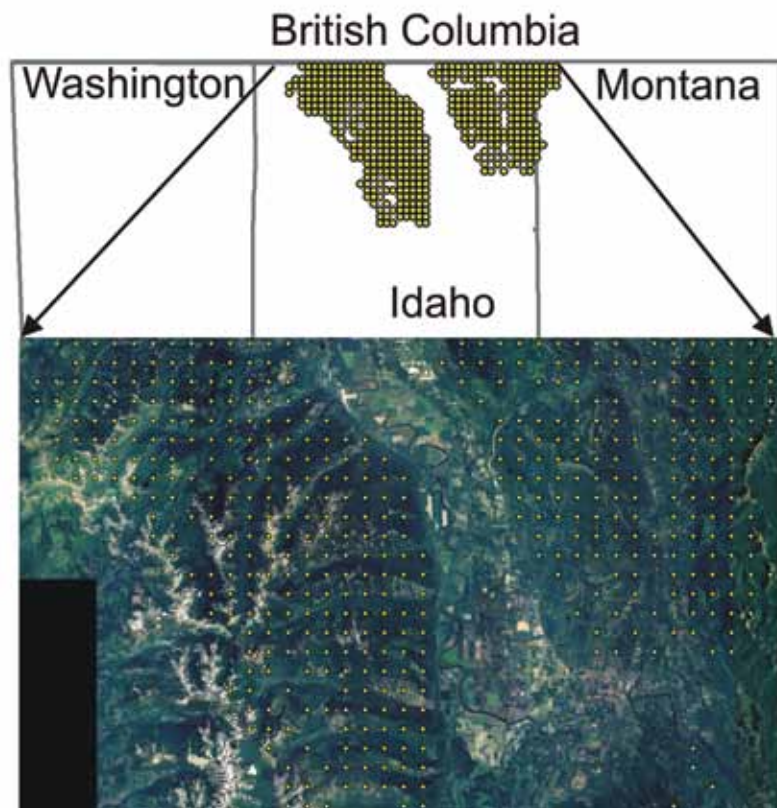


Figure 1. Study area orientation map.

We produced three sites-by-species matrices for the tree data, summarizing the density, basal area, and importance value for each species in each plot. Density is recorded in total number of trees on the plot, basal area is in units of m² per hectare, and importance value is based equally on the relative number of stems and the relative basal area in each plot (Iverson and Prasad 1998). We produced these three versions of the species response data to evaluate the sensitivity of the results to how the dependent variables are measured.

Classified Vegetation Maps

We chose three widely used vegetation maps for the study area to provide a broad evaluation of the range of classified maps currently in use by scientists and managers to represent vegetation community composition and structure at the landscape level. These maps included: (1) Idaho Panhandle National Forest (IPNF) stands map (R1 TSMRS handbook, Art Zack pers. comm.), (2) GAP vegetation type map (<http://gap-analysis.nbi.gov>), and (3) U.S. Forest Service Northern Region Dominance Type Group (R1 VMAP; http://www.fs.fed.us/r1/gis/vmap_v06.html). These maps collectively represent the full range of available spatial

predictions of vegetation assemblages in the study area and are those that are currently being utilized by ecologists and land managers.

The IPNF data consists of photo-interpreted stand delineations (Table 1) and is comparable to stand data currently used by most National Forests for their planning and management decisionmaking (R1 TSMRS handbook, Art Zack pers. comm.). The GAP map was a modification of the Idaho and Montana GAP analysis vegetation type maps. The study area spanned parts of both states, so it was necessary to combine the two maps with reclassification to the patch types shown in Table 1. This map was chosen because GAP maps are often used in landscape-level research and conservation planning. VMAP is a product of the Northern Region U.S. Forest Service and was intended to provide consistent and empirically error-checked vegetation type classification across the 14 national forests in USFS Region 1. It represents “dominance type group,” which is a comparable level of classification resolution as the other maps (Table 1). The three maps collectively represent those typically used in research, conservation, and planning in the northern Rocky Mountains and are representative of the kind and accuracy of such maps used elsewhere.

Table 1. List of classes in the three land cover maps.

Gap code	Gap class Name	IPNF code	IPNF class Name	VMAP Code	VMAP Class name
1	Urban	1	Aspen	3100	Grass
2	Agricultural	2	BGA	3300	Shrub
3	Mining	3	Cedar	5000	Water
6	Water	4	Douglas fir	7000	Sparsely vegetated
7	Rock/soil	5	Grand fir	8010	Ponderosa pine
8	Mixed barren	6	Western larch	8020	Douglas fir
9	Snow	7	Lodgepole pine	8030	Grand fir
10	Clear cut conifer	8	No data	8040	Western larch
11	Alpine meadow	9	Non-forest	8050	Lodgepole pine
12	Burned forest	10	Ponderosa pine	8060	Subalpine fir
13	Forested riparian	11	Subalpine fir	8070	Engelmann spruce
14	Shrub riparian	12	Western hemlock	8090	Western red cedar
15	Grass riparian	13	WLP	8110	Mountain hemlock
16	Wetlands	14	Western white pine	8200	IMXS
19	Grassland	15	WSL	8300	TASH
21	Shrub-grass			8400	TGCH
24	Mesic shrub				
25	Xeric shrub				
26	Aspen				
27	Ponderosa pine				
28	Lodgepole pine				
29	Western Red cedar				
30	Western hemlock				
31	Mixed conifer				
32	Mixed subalpine forest				
33	Mixed whitebark pine forest				
34	Subalpine meadow				

Independent Variables

The independent variable set used in this analysis consisted of the percentage of a 90 m radius window centered on each vegetation plot comprised of each patch type for each of the three classified vegetation maps. There were three motivations for this approach. First, our goal was to evaluate the ability of the classified map to predict species composition of plots across the landscape, so the key variable is the patch type the plot resides in. Second, we wanted to minimize any errors related to patch edge delineation. Classified map error rates tend to be higher near edges of patches where heterogeneity of continuously varying vegetation is highest. The 90 m window size reduces this effect. Third, we chose to compare the ability of the three classified vegetation maps to predict tree community composition and structure and to explicitly compare their independent and over-lapping explanatory ability.

Hypothesis 1: Variance Partitioning With Redundancy Analysis

We used hierarchical variance partitioning (Cushman and McGarigal 2002) with redundancy analysis (RDA) (Legendre and Legendre 1998) to test Hypothesis 1. Hierarchical variance partitioning uses a series of partial canonical ordinations to partition explained variance into its components. The method is described in detail in Cushman and McGarigal (2002). In the analyses in this paper, we completed three separate decompositions, corresponding to the three forms of the dependent variable data set (density, basal area, and importance value by tree species). The decomposition separates the explanatory ability of each classified map. This decomposition extracts seven discrete components of explained community variation and one component of unexplained variation: (1) variance uniquely explained by map IPNF; (2) variance uniquely explained by map GAP; (3) variance uniquely explained by map VMAP; (4) joint explanation of community composition by IPNF and GAP; (5) joint explanation of community composition by GAP and VMAP; (6) joint explanation of community composition by IPNF and VMAP; and (7) joint effects of all three maps.

We conducted each variance decomposition with RDA after row normalization of the species data (dependent variables), as recommended by Legendre and Gallagher (2001). Row normalization converts

the species data to relative profiles within each site. This approach allows for the use of RDA, which is Euclidean-based, for the analysis of community data, while circumventing the problems associated with the Euclidean distance and avoiding certain problems associated with Canonical Correspondence Analysis (CCA) such as the differential weighting of species and sites preserved in the chi-square distance measure of CCA (Legendre and Gallagher 2001). We conducted the variance decomposition both with and without the data standardization using both RDA and CCA and the results did not differ substantially. After standardizing each of the dependent variable sets, we used forward stepwise variable selection based on minimizing Akaike's Information Criterion to select a parsimonious set of variables from each explanatory variable set (i.e., vegetation maps). This resulted in three to nine explanatory variables in the final sets of explanatory variables. All statistical analyses were conducted using the vegan library of R (rproject.org) and scripts written by the authors.

Hypothesis 2: Predicting Occurrence of Four Major Species With Random Forest

Random Forest is a powerful refinement of classification and regression tree (CART) approaches for predicting species occurrence. CART approaches have gained broad usage in ecological studies (Déath and Fabricius 2000). However, CART suffers from several problems, such as over-fitting and difficulty in parameter selection. Several solutions have been proposed that incorporate iterative approaches (Schapire 1990; Breiman 1996). One approach in particular, Random Forests (Breiman 2001), has risen to prominence due to its ability to handle large numbers of predictor variables, accurately reflect complex non-linear species-environment relationships (Cutler and others 2007), and provide local and global measures of variable importance (Evans and Cushman, 2009).

Evans and Cushman (2009) predicted occurrence probabilities for four selected species using the Random Forest algorithmic method (Breiman 2001; Cutler and others 2007) as implemented in R (R development core 2007; Liaw and Wiener 2002). The selected species are *Thuja plicata* (THPL), *Pseudotsuga menziesii* (PSME), *Abies lasiocarpa* (ABLA), and *Pinus ponderosa* (PIPO). These species were selected due to their dominance in each of four major biophysical conditions: THPL dominant in mesic, middle elevation montane

valleys; PSME as a widespread co-dominant on dry to mesic sites at low to middle elevation; ABLA dominant in cold and moist upper elevation sites; and PIPO dominant in warm and dry lower elevations. The analysis predicted occurrence at the species-level and the pixel scale, using the same species occurrence dataset as used in the community-level, stand-scale analysis above and a set of 40 spectral, topographical, and climatic independent variables.

Here we use identical methods to predict the same species, but as functions of the three classified vegetation maps rather than spectral, topographical, and climatic variables. The difference in the analyses is that the independent variable set is the percentage of a 90 m radius window centered on each vegetation plot comprised of each patch type for each of the three classified vegetation maps. The response variable was defined as presence (1) if proportion > 0.10 and absence if less than 0.10. Previous studies (Chawala and others 2002; Chen and others 2004) demonstrated that imbalance between presences and absences bias prediction and model-fit error. To correct this, we used a random down-sample method (Evans and Cushman 2009) to correct for imbalance in the number of presences and absences. We ran 5000 bootstrap replicates (k) with replacement using a 36% out-of-bag (OOB) sample, and we set the number of variables permuted at each node (the m parameter) at $\lceil \text{SQRT}(\text{number of } x \text{ variables}) \rceil$, with a minimum of $m=2$. The final model was built by combining trees from all the resulting random forest models (Evans and Cushman, 2009).

Comparison of these model results with those of Evans and Cushman (2009) provides a test the hypothesis that classified community-level, patch scale vegetation maps will provide similar predictive power as models predicting species-level, pixel scale responses based on continuous environmental variables. We evaluate this (Hypothesis 2) by comparing the kappa statistics and area under the Receiver Operating Characteristic (AUC) curve between the three classified map models and the model based on spectral, topographic, and climatic variables (Evans and Cushman 2009) for all four species. The kappa statistic is a corrected index of classification accuracy that indicates the percentage improvement of the classification over random assignment. The area under the Receiver Operating Characteristic (AUC) curve is a cut-point independent measure of the model's discriminatory ability. AUC values of greater than 0.9 indicate excellent discriminating ability, those between 0.7-0.9 are fair, and values less than 0.7 indicate poor model ability to discriminate between presences and absences.

Results

Hypothesis 1: Variance Partitioning With Canonical Correspondence Analysis

In none of the three decompositions did the three maps jointly explain more than 35% of the variance in the tree community (Table 2, Figure 2). The three maps had the highest explanatory power when the species data were in importance value form, and the lowest when in density form, in which only 16% of species variance was explainable by the three maps in combination (Table 2, Figure 2).

No single vegetation map explained more than 25% of the variance in the tree community in any of the three partitionings (Table 2, Figure 2). In all three analyses, GAP was the poorest predictor. For the density partitioning, GAP explained a total of 6.2% of the variance in the tree community and was able to uniquely explain only 2.1% (after accounting for explanatory ability of the other two maps, Figure 2). Likewise, for the basal area partitioning GAP had total and unique explanatory abilities of 12.2% and 2.9% respectively. In the importance value partitioning, GAP explained 15.5% and 3.6%. The IPNF stands map was the strongest predictor in two of the three partitionings, explaining 22.9% in total and 9.6% uniquely in the basal area analysis and 24.9% total and 8.7% uniquely in the importance value analysis. In contrast, the R1 VMAP map was the strongest predictor in the density partitioning, explaining 9.1% in total and 4.6% uniquely.

Hypothesis 2: Predicting Species Occurrence With Random Forests

Comparing kappa and AUC for four conifer species predicted on the basis of the three classified vegetation maps with those for models based on spectral, topographic, and climatic variables (Evans and Cushman 2009) indicated substantial variability among species in how well they were represented by the classified maps, as well as variability among maps in how well they described the occurrence patterns across species (Table 3). ABLA and PSME were predicted with "fair" accuracy using the classified maps as predictor variables, with AUC values over 0.7 and kappa indicating roughly 50% improvement in classification over random assignment (Table 3). In contrast, THPL had low kappa values and AUC of less than 0.620 and kappa

Table 2. Variance partitioning results for each of the three analyses. There are eight components of variation in each partitioning: VMAP—variance explainable uniquely by the VMAP map; GAP—variance explainable uniquely by the GAP map; IPNF—variance explainable uniquely by the IPNF map; VMAP-GAP—variance jointly explained by VMAP and GAP, but not explained by IPNF; VMAP-IPNF—variance explained jointly by VMAP and IPNF but not explained by GAP; GAP-IPNF—variance explained jointly by GAP and IPNF but not by VMAP; VMAP-GAP-IPNF—variance explained by all three maps. Each component corresponds to a single exclusive (non-overlapping) partition of the total species variance, as can be represented in a Venn diagram (see Figure 2). The remaining “unexplained” variance is usually referred to as the “residual” variance and is equal to the sum of the unconstrained eigenvalues.

Importance Value Partitioning				
	Inertia	Prop. total	Prop. constrained	P value
VMAP	0.215	0.044	0.127	0.020
GAP	0.465	0.036	0.104	0.005
IPNF	0.44	0.087	0.251	0.000
VMAP-GAP	0.039	0.017	0.049	NA
VMAP-IPNF	0.205	0.06.0	0.173	NA
GAP-IPNF	0.061	0.014	0.041	NA
VMAP-GAP-IPNF	0.257	0.088	0.254	NA
Unexplained		0.654		

Basal Area Partitioning				
	Inertia	Prop. total	Prop. constrained	P value
VMAP	0.185	0.036	0.117	0.210
GAP	0.437	0.029	0.094	0.046
IPNF	0.476	0.096	0.313	0.000
VMAP-GAP	0.031	0.012	0.039	NA
VMAP-IPNR	0.182	0.052	0.169	NA
GAP-IPNF	0.044	0.012	0.039	NA
VMAP-GAP-IPNF	0.271	0.069	0.225	NA
Unexplained		0.693		

Density Partitioning				
	Inertia	Prop. total	Prop. constrained	P value
VMAP	0.349	0.046	0.287	0.044
GAP	0.829	0.021	0.131	0.017
IPNF	0.473	0.039	0.244	0.010
VMAP-GAP	0.128	0.008	0.050	NA
VMAP-IPNR	0.123	0.013	0.081	NA
GAP-IPNF	0.022	0.009	0.056	NA
VMAP-GAP-IPNF	0.253	0.024	0.15	NA
Unexplained		0.840		

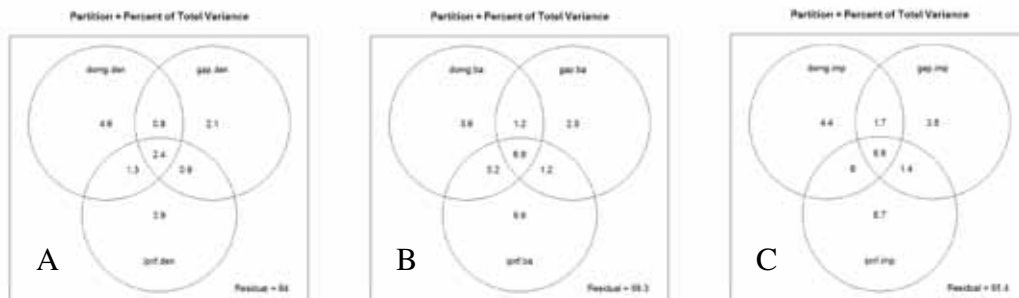


Figure 2. Venn diagrams showing partitioning for the three analyses. A. Density Partitioning, B. Basal Area Partitioning, C. Importance Value Partitioning. Values in the Venn diagram represent proportion of total variance explained.

Table 3. Kappa and AUC for the four study species predicted on the basis of the three classified vegetation maps (IPNF, GAP, VMAP), and the spectral, topographic, and climatic variables utilized in Evans and Cushman (submitted; GRADIENT). The classes included in the three maps are described in Table 1. The kappa statistic is a corrected index of classification accuracy that indicates the percentage the classification is superior to random assignment. The area under the Receiver Operating Characteristic (AUC) curve is a cut-point independent measure of the models discriminatory ability. AUC values of greater than 0.9 indicate excellent discriminating ability, those between 0.7—0.9 are fair, and values less than 0.7 indicate poor model ability to discriminate between presences and absences.

Species	Kappa IPNF	Kappa GAP	Kappa VMAP	Kappa GRADIENT	AUC IPNF	AUC GAP	AUC VMAP	AUC GRADIENT
ABLA	0.559	0.519	0.498	0.97	0.777	0.766	0.749	0.98
PIPO	0.117	-0.005	0.000	0.86	0.538	0.501	0.500	0.98
PSME	0.496	0.333	0.540	0.99	0.731	0.657	0.773	0.99
THPL	0.250	0.083	0.268	0.98	0.612	0.536	0.620	0.99

less than 0.27 for all three maps. Finally, PIPO was very poorly predicted by all three maps, with kappa values less than 0.12 and AUC values very near 0.5 indicating very poor prediction of this species.

In all cases the models developed based on spectral, topographic, and climatic variables (Evans and Cushman 2009) vastly outperformed those based on the classified maps. In all cases, these models showed much higher kappa and AUC values than those produced by any of the classified map based models for any of the four species (Table 3). All four species had kappa and AUC values very near 1 in the Random Forests predictions produced using spectral, topographic, and climatic variables, indicating that the occurrence of these species across the landscape is very highly predictable based on environmental variables. In comparison, the models developed using the classified maps performed much more poorly. The PIPO model completely fails to discriminate presences from absences. The THPL model performs poorly (AUC < 0.620). The models for ABLA and PSME were only marginally effective at discriminating presences from absences, with AUC values over 0.7.

Discussion

Do Classified Vegetation Maps Explain the Composition and Structure of Forest Communities?

For classified vegetation maps to be reliable in representing the structure and composition of the plant community, they must meet at least three critical tests. First, at the community-level they should explain a considerable proportion of the variability in the species distribution data. Second, we would like different classified map products that have the same goal of

representing forest “vegetation types” as patches to provide comparable explanations of community composition (Figure 3). Third, at the species-level, classified maps should be able to explain the variability in the occurrence of individual species nearly as well as models produced directly from selected environmental variables.

Hypothesis 1

In Hypothesis 1a, the question is: Do classified vegetation maps provide a powerful explanation of the structure of the plant community? In Hypothesis 1b, the question is: Do alternative classified vegetation maps provide similar explanations of plant community structure? If the first condition is not met then we have little confidence interpreting the meaning of patch-mosaic maps of vegetation types. If vegetation is not discretely grouped into meaningful associations then there will be no way to produce meaningful spatial classifications into patch mosaic maps. This condition is more fundamental than those typically considered in evaluating classified map accuracy, such as the sufficiency of spectral or textural separability among classes, the adequacy of available imagery or other GIS products to predict class assignment, the availability of sufficient training and truthing data, and identification of the proper grain and extent for analysis. However, if this first condition is not met intrinsically within the ecological system then none of the technical considerations that follow is of importance. If communities do not exist as discrete entities then categorical patch mosaic maps will not meaningfully represent biotic landscape patterns.

Hypothesis 1a: In this case our analysis indicates that the three classified vegetation maps all performed poorly in explaining the composition and structure of

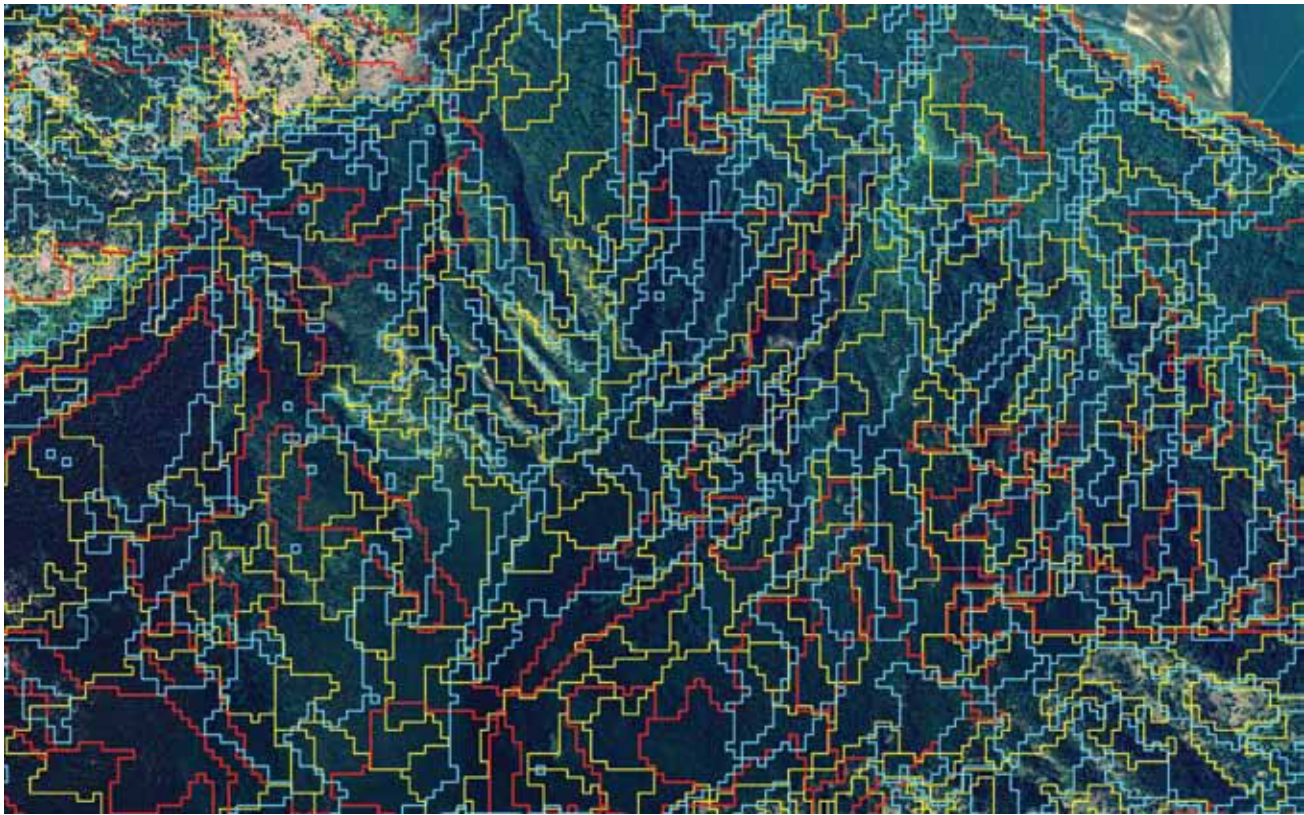


Figure 3. Illustration of incongruity of the three classified vegetation cover type maps with each other and with the underlying pattern of variation in forest vegetation in a portion of the study area. The figure shows the vegetation in a high-resolution color air photo and shows the boundaries of the patches in each of the three classified maps. Blue line—VMAP; Red line—GAP map; Yellow line—Idaho Panhandle National Forest stands map.

forest trees among plots. No map singly was able to explain more than 25% of the variability in the plant community. If over 75% of the variability in species composition in a landscape is unexplainable by a classified map, does that map adequately represent species composition? Recently, Cushman and others (2008) conducted a similar analysis for breeding birds. In that study they found that classified maps of vegetation types were largely unable to predict the occurrence and relative abundance of species within the bird community. This analysis shows that classified vegetation maps also may often fail to predict vegetation community composition itself.

Hypothesis 1b: Even in cases where the first criterion of explanatory power is met, it is important to evaluate whether different map products which purportedly represent the same information provide comparable explanations. If two maps explain largely separate and independent portions of species variability among plots, then they do not provide the same information and cannot be used interchangeably. If this were the case, then the kind of partitioning analysis done in this

paper would be warranted any time a classified map is used to describe vegetation and any time a different map is used to substitute for another. This latter condition is very common, such as when assessments of landscape structure are conducted in different areas with different available map products. It is therefore important to consider how comparable are different maps produced for the same landscape with the same purpose of describing forest vegetation types.

In this analysis the three maps provided largely independent explanations of species variability (Figure 3). For example, in the partitionings based on importance value, basal area, and density, less than 15% of the variability in species data was shared among any combination of two of the maps, and less than 9% was explainable jointly by all three in any of the three cases (Figure 2). These maps were made for the same purpose, to describe the important attributes of vegetation composition across the landscape. That they provide such different and independent information is telling. Two maps made in the same landscape to describe the same ecological attributes at the same spatial scales

should provide comparable information if they are reliable. Indeed, one way to determine the validity of spatial predictions is to look for concordance between independent predictions made using different data. The failure of these maps to provide concordant predictions of vegetation is further evidence that these classified vegetation maps fail to represent the major patterns of vegetation in this landscape (Figure 3).

Hypothesis 2

At the species-level, perhaps the best evaluation of the effectiveness of the classified maps in reflecting the major scales and patterns of variability in the tree community is to compare individual-species models produced based on the classified maps with those produced for the same dependent variable set using topographic, climatic, and spectral independent variables. This allows a formal comparison of model performance based on the two sets of independent variables.

Individual species models based on spectral, topographic, and climatic variables vastly out-performed those produced using the classified maps. The former set of models had kappa and AUC values very near 1 for all four species, indicating very high success in predicting occurrence vs. absence. In contrast, the latter models, in which species occurrence is predicted solely on the basis of the cover types that occur at or in proximity to each vegetation plot, produced much lower kappa and AUC values. Models for one species (PIPO) failed to effectively discriminate presences from absences in any of the three maps, and no models for any species based on any map were able to strongly discriminate between occurrences and absences based on kappa and AUC.

The large differences among species in their relationship to the classified maps further suggests a decoupling of the scales and patterns of individual species response to environmental drivers from the scales and patterns of structure represented in the classified vegetation maps. Further, the high variability in the ability of the three classified maps to predict the four species echoes the point made in evaluating Hypothesis 1b above; different maps made in the same landscape to describe the same ecological attributes at the same spatial scales should provide comparable information. The fact that these maps provide divergent predictions is further evidence that they fail to represent the major patterns of vegetation in this landscape.

Categorical Landscape Ecology as Historical Legacy and Management Convenience

The predominance of classified, categorical maps as the basis for landscape ecology is a historical legacy and a management convenience. German geographer Carl Troll, who introduced the term “landscape ecology” in 1939 (Troll 1971), defined the new field as “the study of the main complex causal relationships between the life communities and their environment in a given section of the landscape.” Troll’s original focus was on continuous patterns of environmental variability and continuous population processes, in a way embracing the gradient concepts of American community ecology proposed by Gleason (1926) and Whittaker (1967). However, landscape ecology shortly thereafter departed from this gradient framework, and evolved into an effort to divide landscapes into small components and ascertain the logic through which they were grouped and interacted as a landscape mosaic (Cushman and others 2010a). This transition was the result of the simultaneous emergence of hierarchy theory and the advent of categorical map analysis in GIS. As is frequently the case, this feedback between theory and methodology led to a “boring in” of the patch mosaic paradigm (Cushman and Huettmann, 2010), such that even as technological advances in remote sensing, GIS and spatial statistics enabled vast improvements in measuring, mapping, and analyzing continuous landscape gradients, the field remained nearly completely dominated by the categorical paradigm (McGarigal and others 2009; Cushman and others 2010b). The dominance of the patch mosaic paradigm was further strengthened by its adoption by land management agencies due to its congruence with ownership maps and the ease of making management decisions based on patch mosaic representations of landscape structure.

Landscape Structure Is Not Vegetation Type

Defining vegetation types is often problematic. Classified vegetation type maps fail to explain a large amount of variance in vegetation structure and composition due to error in map classification, error in delineating categories for the classification, and mismatch between individualistic species responses and categorical definitions. In addition, several other

factors contribute variance to species density, basal area, and importance value, including differences in biophysical setting and seral stage.

It is conceivable that one could stretch the patch-mosaic paradigm to include more of these factors, with unique classes for each vegetation class, in each biophysical setting and each seral stage. However, even if vegetation type classes could be meaningfully and accurately defined, this combination would be unfeasible for several reasons. First, given a modest number of vegetation types, seral stages, and biophysical settings, this combination would result in dozens of individual classes, at a minimum. For example, nine biophysical classes, five seral stages and 10 vegetation types would result in 450 patch types. It is difficult to conceive how to effectively map this many classes with accuracy and even more challenging to analyze and interpret them. Further, it is readily argued that biophysical settings and seral stages themselves are more appropriately represented as gradient phenomena than categorical entities (McGarigal and others 2009; Cushman and others 2010b), which compounds the issues discussed above for categorizing inherently continuous entities.

Toward Gleasonian Landscape Ecology

If classified vegetation maps are poor surrogates for important aspects of vegetation pattern across landscapes, what is the alternative? We believe that explicitly embracing the Gleason-Hutchinson model of individualistic response to spatially structured limiting factors provides a powerful general framework (McGarigal and Cushman 2005; Cushman and others 2007; Evans and Cushman 2009). Specifically, by moving focus from community types to individual species we can avoid the Clementsian pitfall of trying to map and analyze ephemeral and unstable multi-species groups, and instead focus on the realized niches of individual species, which puts analysis within the context of the full scope of ecological theory (Evans and Cushman 2009; Cushman 2010).

This is particularly relevant given the current climate change crisis. Past changes in climate often triggered disassembly of communities. Subsequently, communities reassembled unpredictably, producing mixtures of species that are rare or absent at present (Barnosky and others 1987; Bartlein and others 1997). Typically, displacement of entire vegetation zones or communities did not occur. Instead, species responded individually to climate change according to environmental tolerance, dispersal ability, and biotic interactions (Davis 1986; McLachlan and Brubaker

1995). Because species do not respond en masse, predicting changes in vegetation composition requires species-level resolution; all hierarchical levels above the species are transient.

Accordingly, we advocate addressing vegetation response to climate and disturbance regime at the level of individual species and their interactions with key resources and limiting factors (Cushman and others 2007). Multivariate models characterizing the responses of individual species to gradients of ecological conditions are a primary tool in this effort. These multivariate models predict the response of individual plant species and the composition of plant communities. Importantly, they avoid questionable assumptions about the reality and stability of categorical vegetation types (Gleason 1926; Whittaker 1967).

The fundamental challenge to integrating landscape and community ecology is linking non-spatial niche relationships with the complex patterns of how environmental gradients overlay heterogeneous landscapes (Austin 1985; Cushman 2010). By establishing species optima and tolerances along environmental gradients, researchers can quantify the characteristics of each species' environmental niche. The resulting statistical models can be used to predict the biophysical suitability of each location on a landscape for each species (e.g., Evans and Cushman 2009). This species-level, pixel-scale framework for modeling individualistic species responses can also flexibly include biotic interactions as predictors. This is particularly important, as individualistic species response includes response to the influences of other members of the community as well as abiotic environmental gradients (Lortie and others 2004; Cushman and others 2007). This mapping of niche suitability onto complex landscapes is the fundamental task required to predict individualistic species responses, and it is fundamentally important as a foundation for predicting effects of altered climate and disturbance regimes.

There is no single correct scale of analysis for forest ecology. The fundamental unit of ecological analysis is the organism (Schneider 1994) and fundamental scales are those at which the organism strongly interacts with critical or limiting resources in its environment (Spomer 1973; Cushman and others 2007). Patches, therefore, like "communities," are ephemeral and unstable. Moving from analysis at the "assemblage" to the species level thus simultaneously requires a move from patch to pixel scale. Instead of predicting a categorical entity as a discrete patch, we advocate predicting a continuous entity at each unique location (e.g., Evans and Cushman 2009). These locations become pixels

in the context of GIS and remote sensing databases. Accounting for multiple interactions across ranges of spatial and temporal scales is substantially facilitated by adopting an individualistic, pixel-scale, gradient-based framework (McGarigal and Cushman 2005; Cushman and McKelvey 2010). Studies of relationships between forest communities, climate, and disturbance regimes should use multi-scale approaches that are directly tied to the dominant scales of organism interaction with the environment. Where data allow, it is advantageous to quantitatively measure the relationships among driving factors across a range of scales simultaneously to identify these dominant scales and quantify interaction of factors across scale (e.g., Cushman and McGarigal 2002). Ideally, ecological analysis will therefore not be between hierarchical “levels,” such as populations, communities, or ecosystems, but instead will focus on relationships among organisms and driving processes across continuous ranges of scale (Levin 1992; Cushman 2010; Cushman and others 2010b).

References

- Austin, M. P. 1985. Continuum concept, ordination methods, and niche theory. *Annual Review of Ecology and Systematics*. 16: 39-61.
- Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*. 157: 101-118.
- Barnosky, C. W.; Anderson, P. M.; Bartlein P. J. 1987. The northwestern U.S. during deglaciation: vegetational history and paleoclimatic implications. In: Ruddiman, W. F.; Wright, H. E., Jr., eds. *North America and adjacent oceans during the last deglaciation*. Boulder, CO: Geological Society of America: 289-321.
- Bartlein, P. J.; Whitlock, C.; Shafer, S. 1997. Future climate in the Yellowstone National Park region and its potential impact on vegetation. *Conservation Biology*. 11: 782-792.
- Breiman, L. 1996. Bagging predictors. *Machine Learning*. 24(2): 123-140.
- Brieman, L. 2001. Statistical modeling: the two cultures. *Statistical Science*. 16(3): 199-231.
- Chawala, N. V.; Bowyer, K. W.; Hall, L. O.; Kegelmeyer, W. P. 2002. Smote: synthetic minority oversampling technique. *Journal of Artificial Intelligence Research*. 16: 321-357.
- Chen, C.; Liaw, A.; Breiman, L. 2004. Using random forest to learn imbalanced data. Available: <http://oz.berkeley.edu/users/chenchao/666.pdf>.
- Clements, F. E. 1907. *Plant physiology and ecology*. New York: Hentry Hold & Co.
- Clements, F. E. 1916. *Plant succession: an analysis of the development of vegetation*. Washington, DC: Carnegie Institute of Washington.
- Curtis, J. T.; McIntosh, R. P. 1951. An upland forest continuum in the prairie-forest border region of Wisconsin. *Ecological Monographs*. 32: 476-496.
- Cushman, S. A. 2010. Space and time in ecology: Noise or fundamental driver? In: Cushman, S. A.; Huettman, F., eds. *Spatial complexity, informatics and wildlife conservation*. Tokyo: Springer: 19-42.
- Cushman, S. A.; Huettmann, F. 2010. Ecological knowledge, theory and information in space and time. In: Cushman, S. A.; Huettman, F., eds. *Spatial complexity, informatics and wildlife conservation*. Tokyo: Springer: 3-18.
- Cushman, S. A.; McGarigal, K. 2002. Hierarchical, multiscale decomposition of species-environment relationships. *Landscape Ecology*. 17: 637-646.
- Cushman, S. A.; McKelvey, K. S. 2010. Data on distribution and abundance: monitoring for research and management. In: Cushman, S. A.; Huettman, F., eds. *Spatial complexity, informatics and wildlife conservation*. Tokyo: Springer: 111-130.
- Cushman, S. A.; Evans, J.; McGarigal, K. 2010a. Landscape ecology: past, present and future. In: Cushman, S. A.; Huettman, F., eds. *Spatial complexity, informatics and wildlife conservation*. Tokyo: Springer: 65-82.
- Cushman, S. A.; Littell, J.; McGarigal, K. 2010b. The problem of ecological scaling in spatially complex, non-equilibrium ecological systems. In: Cushman, S. A.; Huettman, F., eds. *Spatial complexity, informatics and wildlife conservation*. Tokyo: Springer: 43-64.
- Cushman, S. A.; McGarigal, K.; Gutzwiller, K.; Evans, J. 2010c. The gradient paradigm: a conceptual and analytical framework for landscape ecology. In: Cushman, S. A.; Huettman, F., eds. *Spatial complexity, informatics and wildlife conservation*. Tokyo: Springer: 83-110.
- Cushman, S. A.; McKelvey, K.; Flather, C.; McGarigal, K. 2008. Testing the use of forest communities to evaluate biological diversity. *Frontiers in Ecology and the Environment*. 6: 13-17.
- Cushman, S. A.; McKenzie, D.; Peterson, D. L.; Littell, J.; McKelvey, K. S. 2007. Research agenda for integrated landscape modelling. Gen. Tech. Rep. RMRS-GTR-194. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 50 p.
- Cutler, D. R.; Edwards, T. C., Jr.; Beard, K. H.; Cutler, A.; Hess, K. T.; Gibson, J.; Lawler, J. 2007. Random forests for classification in ecology. *Ecology*. 88: 2783-2792.
- Davis, M. B. 1986. Climatic instability, time tags, and community disequilibrium. In: Diamond, J.; Case, T., eds. *Community ecology*. New York: Harper and Row: 269-284.
- Déath, G.; Fabricius, K. E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*. 81: 3178-3192.

- Evans, J.; Cushman, S. A. 2009. Gradient modeling of conifer species using random forests. *Landscape Ecology*. 24: 673-683.
- Forman, R. T. T. 1995. *Land mosaics: the ecology of landscapes and regions*. Cambridge: Cambridge University Press.
- Gleason, H. A. 1917. The structure and development of the plant association. *Bulletin of the Torrey Botanical Club*. 43: 463-481.
- Gleason, H. A. 1926. The individualistic concept of the plant association. *Bulletin of the Torrey Botanical Club*. 53: 7-26.
- Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*. 22: 415-427.
- Iverson, L. R.; Prasad, A. M. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs*. 68: 465-485.
- Legendre, P.; Gallagher, E. D. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia*. 129: 271-280.
- Legendre, P.; Legendre, L. 1998. *Numerical ecology. Developments in environmental modelling*, 20. Amsterdam: Elsevier.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology*. 73: 1943-1967.
- Liaw, A.; Wiener, M. 2002. Classification and regression by random forest. *R News: The Newsletter of the R Project*. 2(3): 18-22. Available: <http://cran.r-project.org/doc/Rnews/>.
- Lortie, C. J.; Brooker, R. W.; Choler, P.; [and others]. 2004. Rethinking plant community theory. *Oikos*. 107: 433-438.
- Manning, A. D.; Lindenmayer, D. B.; Nix, H. A. 2004. Continua and umwelt: novel perspectives on viewing landscapes. *Oikos*. 104: 621-628.
- McGarigal, K.; Cushman, S. A. 2005. The gradient concept of landscape structure. In: Wiens, J.; Moss, M., eds. *Issues and perspectives in landscape ecology*. Cambridge: Cambridge University Press: 112-119.
- McGarigal, K.; Tagil, S.; Cushman, S. A. 2009. Surface metrics: an alternative to patch metrics for the quantification of landscape structure. *Landscape Ecology*. 24: 433-450.
- McIntyre, S.; Barrett, G. W. 1992. Habitat variegation, an alternative to fragmentation. *Conservation Biology*. 6: 146-147.
- McLachlan, J. S.; Brubaker, L. B. 1995. Local and regional vegetation change on the northeastern Olympic Peninsula during the Holocene. *Canadian Journal of Botany*. 73: 1618-1627.
- R Development Core Team. 2007. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0. Available: <http://www.R-project.org>.
- Schapire, R. 1990. Strength of weak learnability. *Journal of Machine Learning*. 5: 197-227.
- Schneider, D. C. 1994. *Quantitative ecology: spatial and temporal scaling*. San Diego, CA: Academic Press.
- Spomer, G. G. 1973. The concepts of "interaction" and "operational environment" in environmental analyses. *Ecology*. 54: 200-204.
- ter Braak, C. J. F.; Prentice, I. C. 1988. A theory of gradient analysis. *Advances in Ecological Research*. 18: 271-313.
- Troll, C. 1971. *Landscape Ecology (Geocology) and Biogeocenology—A Terminology Study*. *Geoforum*. 8/71: 43-46.
- Turner, M. G. 1989. Landscape ecology: the effect of pattern on process. *Annual Review of Ecology and Systematics*. 20: 171-197.
- Turner, M. G. 2005. Landscape ecology: what is the state of the science? *Annual Review of Ecology, Evolution, and Systematics*. 36: 319-344.
- Turner, M. G.; Gardner, R. H.; O'Neill, R. V. 2001. *Landscape ecology in theory and practice*. New York: Springer-Verlag.
- Whittaker, R. H. 1967. Gradient analysis of vegetation. *Biological Reviews*. 42: 207-264.



The Rocky Mountain Research Station develops scientific information and technology to improve management, protection, and use of the forests and rangelands. Research is designed to meet the needs of the National Forest managers, Federal and State agencies, public and private organizations, academic institutions, industry, and individuals. Studies accelerate solutions to problems involving ecosystems, range, forests, water, recreation, fire, resource inventory, land reclamation, community sustainability, forest engineering technology, multiple use economics, wildlife and fish habitat, and forest insects and diseases. Studies are conducted cooperatively, and applications may be found worldwide.

Station Headquarters

Rocky Mountain Research Station
240 W Prospect Road
Fort Collins, CO 80526
(970) 498-1100

Research Locations

Flagstaff, Arizona
Fort Collins, Colorado
Boise, Idaho
Moscow, Idaho
Bozeman, Montana
Missoula, Montana

Reno, Nevada
Albuquerque, New Mexico
Rapid City, South Dakota
Logan, Utah
Ogden, Utah
Provo, Utah

The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD). To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, DC 20250-9410, or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.

